Linear MMSE Precoder Combiner Designs for Decentralized Estimation in Wireless Sensor Networks

Kunwar Pritiraj Rajput[†], Yogesh Verma*, Naveen K. D. Venkategowda**, Aditya K. Jagannatham[†] and Pramod K. Varshney[‡]
[†]Department of Electrical Engineering, Indian Institute of Technology Kanpur, India
*Intel Technology India Pvt. Ltd., Bangalore, India
**Department of Electronic Systems, Norwegian University of Science and Technology, Trondheim, Norway
[‡]Department of Electrical Engineering and Computer Science, Syracuse University, NY, USA
(e-mail: pratiraj@iitk.ac.in, yogesh.verma@intel.com, naveen.dv@ntnu.no, adityaj@iitk.ac.in, varshney@syr.edu)

***Abstract*—This work considers the design of linear minimum mean square error (MMSE) precoders and combiners for the estimation of an unknown vector parameter in a coherent multiple access channel (MAC)-based multiple-input multiple-output (MIMO) wireless sensor network. The proposed designs that minimize the mean squared error (MSE) of the parameter estimate at the fusion center are based on majorization theory, which leads to non-iterative closed-form solutions for the precoders and combiners. Various scenarios are considered for parameter estimation such as networks with ideal high precision sensors as well as noisy non-ideal sensors. Moreover, inter parameter correlation is also incorporated, which makes the analysis comprehensive. The Bayesian Cramer-Rao bound (BCRB) and centralized MMSE bound are determined to characterize the estimation performance. Simulation results demonstrate the improved performance and also corroborate our analytical formulations.**

***Index Terms*—Wireless sensor networks, precoder-combiner design, linear decentralized estimation, majorization theory.**

I. Introduction

The advances in wireless communication technologies, coupled with the lower fabrication cost of smart sensors, have led to the development of large networks of miniature sensor nodes connected over wireless links, termed wireless sensor networks (WSNs) [1]. More recently, this has enabled the deployment of the Internet of Things (IoT) that integrates a large number of sensor equipped devices with the 5G network. A commonly used WSN comprises of a group of spatially co-located sensors that continuously sense physical phenomena, followed by the transmission of suitably pre-processed data to the fusion center for subsequent signal processing and analysis. Since the sensors often are bandwidth and power constrained, it is imperative to design optimal pre-processing or precoding strategies at the sensors for efficient transmission of the observations, and also efficient post-processing or combining techniques at the fusion center to maximize the accuracy of parameter estimation. Furthermore, multiple-input multiple-output (MIMO) technology that enables the simultaneous transmission of multiple observations, can play a significant role in enhancing the bandwidth efficiency of the sensor network. Naturally, the design of precoding and combining schemes for parameter estimation in MIMO sensor networks has attracted significant research interest. A brief review of related works in the existing literature is presented next.

A. *Review of existing works*

The pioneering work in [2] first introduced a general model for the decentralized estimation of scalar as well as vector parameters in a sensor network. For transmission of the measurements from the sensors to the fusion center (FC), the research therein considered an orthogonal MAC, in which the sensor transmissions are over parallel orthogonal channels, as well as a coherent (additive) MAC, in which all the transmissions are over a single channel. Thus, the transmissions in the latter channel are superimposed at the fusion center. The authors demonstrated the fundamental result that the coherent MAC, with optimal power allocation, can lead to a similar performance of decentralized estimation as that of an orthogonal MAC. This renders the former scheme bandwidth efficient, while not compromising on the performance of parameter estimation. However, to simplify the pertinent analysis, the authors in [2] assumed the channel matrix between each sensor and the fusion center to be diagonal in nature. This restricts the applicability of the framework proposed therein. To overcome this shortcoming, iterative algorithms were developed using the block coordinate descent framework in [3] and [4], by considering a general non-diagonal MIMO channel matrix between each sensor and the fusion center. However, due to their iterative nature, the algorithms described therein require significant computational resources and time. In [5], the authors present precoder design that eliminates the need of a combiner, but such design can potentially lead to noise enhancement due to the zero-forcing nature of the distortionless constraint. Authors in [6], [7] propose schemes for the sequential LMMSE estimation of time varying scalar and vector parameters, respectively. The algorithms developed therein are once again iterative in nature. The problem of linear decentralized estimation has been extended to massive MIMO 5G systems in [8], with a very large antenna array at the fusion center and only single antenna sensor nodes. In a paradigm shift, sensor collaboration schemes wherein the constituent sensors exchange and pre-process their observations prior to transmission in order to reduce the communication cost and improve estimation efficiency are presented in [9]–[11]. However, they can lead to a significant communication overhead due to the inter-sensor communication that is necessitated by such schemes. To overcome the various shortcomings of the works reviewed above, this paper develops linear processing schemes for decentralized vector parameter estimation in a MIMO sensor network.

B. Contributions of the work

This paper considers a MIMO WSN with multiple antennas at the sensor nodes as well as at the fusion center for the transmission and estimation of a vector parameter. The proposed linear MMSE processing techniques are based on majorization theory [12], which leads to closed form solutions for the design of precoders and combiners, unlike the iterative schemes in the existing works [3], [4].

- The proposed precoding/ combining strategies to minimize the MSE of parameter estimation at the fusion center under different power constraints are first developed for a general scenario with ideal noiseless sensors.
- These are subsequently extended to a more general scenario with non-ideal noisy sensors as well as parameter correlation. The design for the case of an uncorrelated white parameter process is presented as a special case.
- The BCRB and centralized MMSE bounds are also determined to benchmark the performance of the proposed communication and estimation paradigms for the scenarios with ideal and non-ideal sensors, respectively.
- Simulation results demonstrate the efficacy of the proposed designs and a performance close to the benchmarks.

The remainder of the paper is organized as follows: Section II describes the MIMO WSN system model, while Section III presents linear precoder and combiner designs that minimize the MSE for scenario with ideal high precision sensors. In Section IV precoder and combiner designs are developed for a general scenario with non-ideal noisy sensor observations, first for the general case of correlated parameters followed by the special case of uncorrelated parameter estimation. Simulation results are provided in Section V, followed by the conclusion in Section VI and proof of the optimality of the proposed precoder design is given in the Appendix A.

II. System Model

Consider L wireless sensors communicating with the fusion center over a coherent MAC [2]. Let N_t and N_r denote the number of antennas at each sensor node and the fusion center, respectively. The observation vector $\mathbf{x}_i \in \mathbb{C}^{l_i \times 1}$ of the ith sensor can be modeled as

$$\mathbf{x}_i = \mathbf{G}_i \boldsymbol{\theta} + \mathbf{v}_i, \tag{1}$$

where $\boldsymbol{\theta} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_\theta)$ is the m dimensional parameter vector of interest, $\mathbf{G}_i \in \mathbb{C}^{l_i \times m}$ represents the corresponding observation matrix, and $\mathbf{v}_i \in \mathbb{C}^{l_i \times 1}$ denotes the observation noise that is distributed as $\mathcal{CN}(\mathbf{0}, \mathbf{R}_i)$. The observation vector $\mathbf{x}_i$ at sensor i is initially processed employing the whitening filter $\mathbf{W}_i \in \mathbb{C}^{l_i \times l_i}$ to yield the output $\mathbf{W}_i \mathbf{x}_i = \mathbf{W}_i \mathbf{G}_i \boldsymbol{\theta} + \mathbf{W}_i \mathbf{v}_i$. Subsequently, the whitened vector is precoded with the matrix $\mathbf{B}_i \in \mathbb{C}^{N_t \times l_i}$, prior to transmission. Let $\mathbf{H}_i \in \mathbb{C}^{N_r \times N_t}$ denote the channel between the ith sensor and the fusion center. Hence, the received vector $\mathbf{y} \in \mathbb{C}^{N_r \times 1}$at the fusion center over the coherent MAC can be modeled as

$$\begin{aligned}\mathbf{y} &= \sum_{i=1}^{L} \mathbf{H}_i \mathbf{B}_i \mathbf{W}_i \mathbf{G}_i \boldsymbol{\theta} + \sum_{i=1}^{L} \mathbf{H}_i \mathbf{B}_i \mathbf{W}_i \mathbf{v}_i + \mathbf{n}, \\ &= \mathbf{HBWG}\boldsymbol{\theta} + \mathbf{HBWv} + \mathbf{n},\end{aligned} \tag{2}$$

where the quantities $\mathbf{H} \in \mathbb{C}^{N_r \times LN_t}$, $\mathbf{W} \in \mathbb{C}^{l \times l}$, $\mathbf{B} \in \mathbb{C}^{LN_t \times l}$ and $\mathbf{R}_v \in \mathbb{C}^{l \times l}$, denote the concatenated channel matrix $\mathbf{H} = [\mathbf{H}_1, \mathbf{H}_2, \ldots, \mathbf{H}_L]$, block diagonal whitening matrix $\mathbf{W} = \operatorname{diag}(\mathbf{W}_1, \mathbf{W}_2, \ldots, \mathbf{W}_L)$, block diagonal precoding matrix $\mathbf{B} = \operatorname{diag}(\mathbf{B}_1, \mathbf{B}_2, \ldots, \mathbf{B}_L)$, and observation noise covariance matrix $\mathbf{R}_v = \operatorname{diag}(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_L)$, respectively. The stacked observation matrix $\mathbf{G} \in \mathbb{C}^{l \times m}$ for the network is defined as $\mathbf{G} = [\mathbf{G}_1^H, \mathbf{G}_2^H, \ldots, \mathbf{G}_L^H]^H$. Furthermore, $\mathbf{v} = [\mathbf{v}_1^H, \mathbf{v}_2^H, \ldots, \mathbf{v}_L^H]^H \in \mathbb{C}^{l \times 1}$ denotes the stacked observation noise vector, $\mathbf{n} \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_n)$ represents the N_r dimensional receiver noise and $l = \sum_{i=1}^{L} l_i$.

Subsequently, the linear minimum mean square error (LMMSE) combiner is used to generate the estimate $\hat{\boldsymbol{\theta}}$ of the parameter vector $\boldsymbol{\theta}$ at the fusion center. The error covariance matrix $\mathbf{E} \in \mathbb{C}^{m \times m}$ of the estimate $\hat{\boldsymbol{\theta}}$ is defined as

$$\mathbf{E} = \mathbb{E}\left\{ \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right) \left(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}\right)^H \right\}. \tag{3}$$

In order to constrain the transmit power of the sensor network, the transmit power of the ith sensor can be derived as

$$\begin{aligned}&\mathbb{E}\left\{ ||\mathbf{B}_i \mathbf{x}_i||_2^2 \right\} = \operatorname{Tr}\Big(\mathbf{B}_i \mathbb{E}\left\{ \mathbf{x}_i \mathbf{x}_i^H \right\}) \mathbf{B}_i^H \Big) \\ &\quad = \operatorname{Tr}\Big(\mathbf{B}_i \left(\mathbf{W}_i \mathbf{G}_i \mathbf{R}_\theta \mathbf{G}_i^H \mathbf{W}_i^H + \mathbf{W}_i \mathbf{R}_i \mathbf{W}_i^H \right) \mathbf{B}_i{}^H \Big).\end{aligned} \tag{4}$$

From (4), the total transmit power of the WSN can be evaluated as

$$\begin{aligned}&\sum_{i=1}^{L} \mathbb{E}\left\{ ||\mathbf{B}_i \mathbf{x}_i||_2^2 \right\} \\ &= \operatorname{Tr}\Big(\mathbf{B} \left(\mathbf{WGR}_\theta \mathbf{G}^H \mathbf{W}^H + \mathbf{WR}_v \mathbf{W}^H \right) \mathbf{B}^H \Big).\end{aligned} \tag{5}$$

From (3) and (5), the optimization problem for the design of the sensor precoding matrices $\mathbf{B}_i$, $i = 1, 2, \ldots, L$, and the combiner at the fusion center, to minimize the MSE of estimation with the total power constrained to P_T, can be formulated as

$$\begin{aligned}\min_{\{\mathbf{B}\}_{i=1}^{L}} \quad & \operatorname{Tr}(\mathbf{E}) \\ \text{s.t.} \quad & \operatorname{Tr}\Big(\mathbf{B} \left(\mathbf{WGR}_\theta \mathbf{G}^H \mathbf{W}^H + \mathbf{WR}_v \mathbf{W}^H \right) \mathbf{B}^H \Big) \le P_T.\end{aligned} \tag{6}$$

Majorization theory is employed next to determine precoder and combiner designs for different scenarios in the sensor network.

III. Precoder Combiner Design with Ideal Sensors

To begin with, this section considers a scenario with ideal high-precision sensors where the observation noise can be ignored. A more general scenario with non-ideal noisy sensors, and also incorporating correlation among the phenomena being sensed, is analyzed in later sections. The received signal $\mathbf{y}$ in (2) with ideal sensors is given by

$$\mathbf{y} = \mathbf{H}\mathbf{B}\mathbf{G}\boldsymbol{\theta} + \mathbf{n} = \mathbf{H}\mathbf{F}\boldsymbol{\theta} + \mathbf{n}, \tag{7}$$

where the matrix $\mathbf{F} \in \mathbb{C}^{LN_t \times m}$ is defined as $\mathbf{F} = \mathbf{B}\mathbf{G} = [\mathbf{F}_1^H, \mathbf{F}_2^H, \ldots, \mathbf{F}_L^H]^H$ and $\mathbf{F}_i = \mathbf{B}_i\mathbf{G}_i \in \mathbb{C}^{N_t \times l_i}$. Subsequently, employing the LMMSE combiner $\mathbf{A} = \left(\mathbf{H}\mathbf{F}\mathbf{R}_\theta\mathbf{F}^H\mathbf{H}^H + \mathbf{R}_n\right)^{-1}\mathbf{H}\mathbf{F}\mathbf{R}_\theta \in \mathbb{C}^{N_r \times m}$, the LMMSE estimate is obtained as $\hat{\boldsymbol{\theta}} = \mathbf{A}^H\mathbf{y}$. The corresponding error covariance matrix $\mathbf{E}$ and the resulting MSE are obtained as [13]

$$\mathbf{E} = \left(\mathbf{R}_\theta^{-1} + \mathbf{F}^H\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F}\right)^{-1} \tag{8}$$

$$\text{MSE} = \text{Tr}\left(\mathbf{E}\right) = \text{Tr}\left(\left(\mathbf{R}_\theta^{-1} + \mathbf{F}^H\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F}\right)^{-1}\right). \tag{9}$$

The transmit power constraint reduces to $\text{Tr}\left(\mathbf{F}\mathbf{R}_\theta\mathbf{F}^H\right) \leq P_T$. Hence, the optimization problem in (6) for MSE minimization reduces to

$$\begin{aligned} \underset{\mathbf{F}}{\text{minimize}} \quad & \text{Tr}\left(\left(\mathbf{R}_\theta^{-1} + \mathbf{F}^H\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}\mathbf{F}\right)^{-1}\right) \\ \text{subject to} \quad & \text{Tr}\left(\mathbf{F}\mathbf{R}_\theta\mathbf{F}^H\right) \leq P_T. \end{aligned} \tag{10}$$

Majorization theory [12] presents a mathematically tractable framework for precoder design via diagonalization of the error covariance matrix. Substituting the eigenvalue decompositions of the different matrices as $\mathbf{R}_\theta$=$\mathbf{U}_\theta\boldsymbol{\Lambda}_\theta\mathbf{U}_\theta^H$ and $\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H$ in (9) above, one obtains the equivalent expression for the MSE as

$$\text{MSE} = \text{Tr}\Bigg(\mathbf{U}_\theta \underbrace{\left(\boldsymbol{\Lambda}_\theta^{-1} + \mathbf{U}_\theta^H\mathbf{F}^H\mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^H\mathbf{F}\mathbf{U}_\theta\right)^{-1}}_{\bar{\mathbf{E}}} \mathbf{U}_\theta^H\Bigg). \tag{11}$$

Let the matrix $\mathbf{F}$ be chosen to have the structure $\mathbf{F} = \mathbf{Q}\boldsymbol{\Omega}\mathbf{U}_\theta^H$, where $\boldsymbol{\Omega}$ is the $LN_t \times m$ matrix given by

$$\boldsymbol{\Omega} = \begin{bmatrix} \left(\text{diag}\left(\mathbf{p}\right)\right)_{m\times m}^{\frac{1}{2}} \\ \mathbf{0}_{(LN_t - m)\times m} \end{bmatrix}, \tag{12}$$

where the vector $\mathbf{p} = [p(1), p(2), \cdots, p(m)]^T \in \mathbb{R}^{m\times 1}$, $p(j)$ denoting the jth element of $\mathbf{p}$ and $\mathbf{0}_{(LN_t - m)\times m}$ denotes a $(LN_t - m) \times m$ matrix with all its elements equal to zero. Furthermore, $\text{diag}\left(\mathbf{p}\right)$ denotes a $m \times m$ diagonal matrix, with the elements of vector $\mathbf{p}$ on its principal diagonal. It is demonstrated in Appendix A that such a precoder structure, which diagonalizes the effective error covariance matrix $\bar{\mathbf{E}}$ above, is optimal due to the Schur concave nature of the MSE cost function [12]. Note that the MSEs corresponding to $\mathbf{E}$ and $\bar{\mathbf{E}}$ are identical since $\text{Tr}\left(\mathbf{E}\right) = \text{Tr}(\bar{\mathbf{E}})$. Employing the above substitution, the resulting MSE can be simplified as

$$\begin{aligned} \text{MSE} &= \text{Tr}\left(\left(\boldsymbol{\Lambda}_\theta^{-1} + \boldsymbol{\Omega}^H\boldsymbol{\Lambda}\boldsymbol{\Omega}\right)^{-1}\right) \\ &= \sum_{j=1}^{m} \frac{1}{\frac{1}{\lambda_j(\mathbf{R}_\theta)} + p(j)\lambda_j(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H})}, \end{aligned} \tag{13}$$

where the quantities $\lambda_j(\mathbf{R}_\theta)$ and $\lambda_j(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H})$ denote the jth eigenvalue of the matrices $\mathbf{R}_\theta$ and $\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H}$, respectively. Further, the total transmit power constraint in (10) can be simplified as

$$\text{Tr}\left(\mathbf{F}\mathbf{R}_\theta\mathbf{F}^H\right) = \text{Tr}\left(\boldsymbol{\Omega}\boldsymbol{\Lambda}_\theta\boldsymbol{\Omega}^H\right) = \sum_{j=1}^{m} p(j)\lambda_j(\mathbf{R}_\theta). \tag{14}$$

Hence, form (13) and (14) the optimization problem in (10) to minimize the MSE can be equivalently formulated as

$$\begin{aligned} \underset{\mathbf{p}}{\text{minimize}} \quad & \sum_{j=1}^{m} \frac{1}{\frac{1}{\lambda_j(\mathbf{R}_\theta)} + p(j)\lambda_j(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H})} \\ \text{subject to} \quad & \sum_{j=1}^{m} p(j)\lambda_j(\mathbf{R}_\theta) \leq P_T \\ & p(j) \geq 0, \ j = 1, 2 \ldots, m. \end{aligned} \tag{15}$$

Using the Karush-Kuhn-Tucker (KKT) framework [14], the optimal values of $p(j)$ are derived as

$$p(j) = \left(\mu\sqrt{\frac{1}{\lambda_j(\mathbf{R}_\theta)\lambda_j(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H})}} - \frac{1}{\lambda_j(\mathbf{R}_\theta)\lambda_j(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H})}\right)^+, \tag{16}$$

and the Lagrange multiplier μ that satisfies the power constraint with equality is determined as

$$\mu = \frac{P_T + \sum_{j=1}^{m} \frac{1}{\lambda_j(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H})}}{\sum_{j=1}^{m} \sqrt{\frac{\lambda_j(\mathbf{R}_\theta)}{\lambda_j(\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{H})}}}. \tag{17}$$

The optimal values $p(j)$ upon substitution in the expression for $\boldsymbol{\Omega}$ in (12) yield the matrix $\mathbf{F}$. The individual precoders $\mathbf{B}_i$ can be determined as $\mathbf{B}_i = \mathbf{F}_i\mathbf{G}_i^\dagger$, where $\mathbf{G}_i^\dagger$ denotes the pseudo-inverse of $\mathbf{G}_i$. One can also solve the MSE minimization problem with individual sensor power constraints as follows. Let P_i denote the maximum power of the ith sensor. The power constraint for the ith sensor in (4) reduces to $\text{Tr}\left(\mathbf{F}_i\mathbf{R}_\theta\mathbf{F}_i^H\right) \leq P_i$. Let $\mathbf{F}_i = \mathbf{Q}_i\boldsymbol{\Omega}\mathbf{U}_\theta^H$, where $\mathbf{Q}_i$ denotes the sub-matrix of $\mathbf{Q}$ corresponding to rows $(i-1)N_t + 1$ to iN_t and all the columns. The power of the ith sensor is $\sum_{j=1}^{m} p(j)\lambda_j(\mathbf{R}_\theta)[\mathbf{Q}_i^H\mathbf{Q}_i]_{jj}$. Replacing the total power budget in (15) with L power constraints corresponding to the individual sensors, one can determine the optimal precoders for this scenario using convex solvers such as CVX [15].

$$\text{MSE} = \text{Tr}\left(\left(\mathbf{V}_{\tilde{\mathbf{G}}}^H \mathbf{R}_\theta^{-1} \mathbf{V}_{\tilde{\mathbf{G}}} + \mathbf{\Sigma}_{\tilde{\mathbf{G}}}^H \mathbf{U}_{\tilde{\mathbf{G}}}^H \tilde{\mathbf{F}}^H \mathbf{U}_n \left(\mathbf{U}_n^H \tilde{\mathbf{F}} \tilde{\mathbf{F}}^H \mathbf{U}_n + \mathbf{\Lambda}_n\right)^{-1} \mathbf{U}_n^H \tilde{\mathbf{F}} \mathbf{U}_{\tilde{\mathbf{G}}} \mathbf{\Sigma}_{\tilde{\mathbf{G}}}\right)^{-1}\right). \quad (19)$$

IV. Precoder Combiner Design with Non-ideal Sensors

This section now considers precoder design and parameter estimation at the fusion center for a general scenario with non-ideal noisy sensors, and also incorporates parameter correlation. As described in the system model in Section II, the LMMSE combiner $\mathbf{A} = \left(\tilde{\mathbf{F}}\tilde{\mathbf{G}}\mathbf{R}_\theta\tilde{\mathbf{G}}^H\tilde{\mathbf{F}}^H + \tilde{\mathbf{F}}\tilde{\mathbf{F}}^H + \mathbf{R}_n\right)^{-1}\tilde{\mathbf{F}}\tilde{\mathbf{G}}\mathbf{R}_\theta \in \mathbb{C}^{N_r \times m}$, is used to generate the estimate of the underlying parameter $\boldsymbol{\theta}$ from the received vector $\mathbf{y}$ in (2). The resulting MSE is given as [13]

$$\text{MSE} = \text{Tr}\left(\left(\mathbf{R}_\theta^{-1} + \tilde{\mathbf{G}}^H\tilde{\mathbf{F}}^H\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H + \mathbf{R}_n\right)^{-1}\tilde{\mathbf{F}}\tilde{\mathbf{G}}\right)^{-1}\right), \quad (18)$$

where the matrices $\tilde{\mathbf{F}} = \mathbf{H}\mathbf{B} = [\tilde{\mathbf{F}}_1, \tilde{\mathbf{F}}_2, \ldots, \tilde{\mathbf{F}}_L] \in \mathbb{C}^{N_r \times l}$, $\tilde{\mathbf{G}} = \mathbf{W}\mathbf{G} = [(\mathbf{W}_1\mathbf{G}_1)^T (\mathbf{W}_2\mathbf{G}_2)^T, \ldots, (\mathbf{W}_L\mathbf{G}_L)^T]^T \in \mathbb{C}^{l \times m}$ and $\tilde{\mathbf{F}}_i = \mathbf{H}_i\mathbf{B}_i \in \mathbb{C}^{N_r \times l_i}$.

Let the singular value decomposition of $\tilde{\mathbf{G}} = \mathbf{U}_{\tilde{G}}\mathbf{\Sigma}_{\tilde{G}}\mathbf{V}_{\tilde{G}}^H$ and the eigenvalue decomposition of $\mathbf{R}_n = \mathbf{U}_n\mathbf{\Lambda}_n\mathbf{U}_n^H$. Substituting these in the expression for the MSE in (18), the resulting expression can be simplified to the one shown in (19). Defining $\mathbf{\Gamma} = \mathbf{V}_{\tilde{G}}^H\mathbf{R}_\theta^{-1}\mathbf{V}_{\tilde{G}}$ and setting the precoder $\tilde{\mathbf{F}} = \mathbf{U}_n\tilde{\mathbf{\Omega}}\mathbf{U}_{\tilde{G}}^H$, where $\tilde{\mathbf{\Omega}} \in \mathbb{C}^{N_r \times l}$ is defined as

$$\tilde{\mathbf{\Omega}} = \begin{bmatrix} (\text{diag}(\mathbf{p}))_{m\times m}^{\frac{1}{2}} \, \mathbf{0}_{m\times(l-m)} \\ \mathbf{0}_{(N_r-m)\times l} \end{bmatrix}, \quad (20)$$

the MSE in (19) becomes

$$\text{MSE} = \text{Tr}\left((\mathbf{\Gamma} + \mathbf{D})^{-1}\right), \quad (21)$$

where $\mathbf{D} = \mathbf{\Sigma}_{\tilde{G}}^H\tilde{\mathbf{\Omega}}^H\left(\tilde{\mathbf{\Omega}}\tilde{\mathbf{\Omega}}^H + \mathbf{\Lambda}_n\right)^{-1}\tilde{\mathbf{\Omega}}\mathbf{\Sigma}_{\tilde{G}}$. Since the matrix $\mathbf{\Gamma}$ is not necessarily diagonal, the following result from [16], is employed to simplify the precoder design problem.

Theorem 1 (Weyl's Theorem)**.** *Let* $\mathbf{X}$, $\mathbf{Y} \in \mathbb{C}^{n\times n}$ *denote Hermitian symmetric matrices. Further, let the eigenvalues of* $\mathbf{X}$, $\mathbf{Y}$ *and* $\mathbf{X}+\mathbf{Y}$ *be arranged in increasing order. Then*

$$\lambda_{j+k-1}(\mathbf{X}+\mathbf{Y}) \geq \lambda_k(\mathbf{X}) + \lambda_j(\mathbf{Y}),$$

for every pair of integers j, k *such that* $1 \leq j, k \leq n$ *and* $j + k \leq n + 1$.

Setting $k = 1$ in the above result, one can determine an upper bound on the MSE as

$$\text{MSE} = \text{Tr}\left((\mathbf{\Gamma}+\mathbf{D})^{-1}\right) \leq \sum_{j=1}^{m}\left(\frac{1}{\lambda_1(\mathbf{\Gamma}) + \frac{p(j)\sigma_j^2(\tilde{\mathbf{G}})}{p(j)+\lambda_i(\mathbf{R}_n)}}\right)$$
$$= \sum_{j=1}^{m}\left(\frac{p(j)+\lambda_j(\mathbf{R}_n)}{p(j)(\lambda_1(\mathbf{\Gamma})+\sigma_j^2(\tilde{\mathbf{G}})) + \lambda_1(\mathbf{\Gamma})\lambda_j(\mathbf{R}_n)}\right), \quad (22)$$

where $\sigma_j\left(\tilde{\mathbf{G}}\right)$ is the jth diagonal element of the singular value matrix $\mathbf{\Sigma}_{\tilde{G}}$. The total transmit power in (5), after substituting $\mathbf{B}_i = \mathbf{H}_i^\dagger\tilde{\mathbf{F}}_i = \mathbf{H}_i^\dagger\mathbf{U}_n\tilde{\mathbf{\Omega}}\mathbf{U}_{\tilde{\mathbf{G}}_i}^H$ and using the property $\text{Tr}(\mathbf{XY}) = \text{Tr}(\mathbf{YX})$, can be simplified as

$$\sum_{i=1}^{L}\text{Tr}\left(\mathbf{B}_i\left(\mathbf{W}_i\mathbf{G}_i\mathbf{R}_\theta\mathbf{G}_i^H\mathbf{W}_i^H + \mathbf{I}_{l_i}\right)\mathbf{B}_i{}^H\right)$$
$$= \sum_{i=1}^{L}\text{Tr}\left(\mathbf{\Phi}_i\tilde{\mathbf{\Omega}}\mathbf{\Psi}_i\tilde{\mathbf{\Omega}}^H\right) \leq \sum_{i=1}^{L}\text{Tr}(\mathbf{\Psi}_i)\sum_{j=1}^{m}p(j)[\mathbf{\Phi}_i]_{jj}, \quad (23)$$

where the matrices $\mathbf{\Phi}_i = \mathbf{U}_n^H\left(\mathbf{H}_i^\dagger\right)^H\mathbf{H}_i^\dagger\mathbf{U}_n$, $\mathbf{\Psi}_i = \mathbf{U}_{\tilde{\mathbf{G}}_i}^H\mathbf{U}_{\tilde{\mathbf{G}}_i}\left(\mathbf{\Sigma}_{\tilde{\mathbf{G}}}\mathbf{V}_{\tilde{\mathbf{G}}}^H\mathbf{R}_\theta\mathbf{V}_{\tilde{\mathbf{G}}}\mathbf{\Sigma}_{\tilde{G}}^H + \mathbf{I}_{l_i}\right)\mathbf{U}_{\tilde{\mathbf{G}}_i}^H\mathbf{U}_{\tilde{\mathbf{G}}_i}$, and $\mathbf{U}_{\tilde{\mathbf{G}}_i}$ is a sub-matrix of $\mathbf{U}_{\tilde{\mathbf{G}}}$ corresponding to the rows from $\sum_{i=1}^{i-1}l_i+1$ to $\sum_{i=1}^{i}l_i$ and all the columns. The last step of the above simplification is obtained using the property $\text{Tr}(\mathbf{XY}) \leq \text{Tr}(\mathbf{X})\,\text{Tr}(\mathbf{Y})$ [16]. Therefore, the optimization problem for MSE minimization can now be formulated using the results in (22) and (23) as follows

$$\begin{aligned} \underset{\mathbf{p}}{\text{minimize}} \quad & \sum_{j=1}^{m}\frac{p(j)+\lambda_j(\mathbf{R}_n)}{p(j)(\lambda_1(\mathbf{\Gamma})+\sigma_j^2(\tilde{\mathbf{G}}))+\lambda_1(\mathbf{\Gamma})\lambda_j(\mathbf{R}_n)} \\ \text{subject to} \quad & \sum_{i=1}^{L}\text{Tr}(\mathbf{\Psi}_i)\sum_{j=1}^{m}p(j)[\mathbf{\Phi}_i]_{jj} \leq P_T \\ & p(j) \geq 0,\ j = 1, 2\ldots, m. \end{aligned} \quad (24)$$

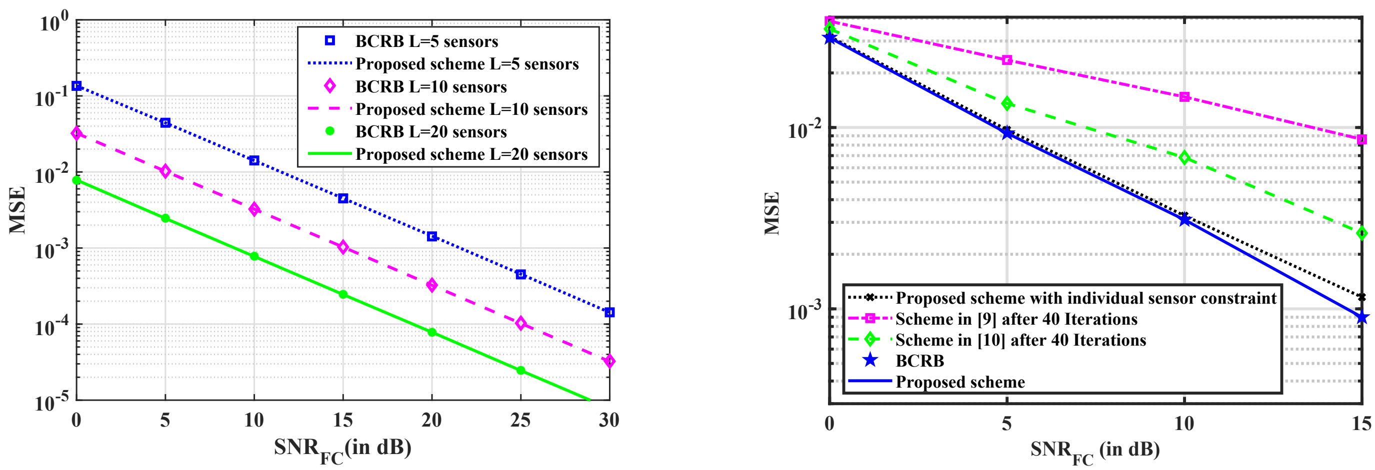
The above optimization problem can once again be solved using the KKT framework [14] to determine the optimal values of $p(j)$ as

$$p(j) = \frac{1}{\alpha_j}\left(\tilde{\mu}\sqrt{\frac{\sigma_j^2(\tilde{\mathbf{G}})\lambda_j(\mathbf{R}_n)}{\sum_{i=1}^{L}\text{Tr}(\mathbf{\Psi}_i)[\mathbf{\Phi}_i]_{jj}}} - \lambda_1(\mathbf{\Gamma})\lambda_j(\mathbf{R}_n)\right)^+, \quad (25)$$

where $\alpha_j = \left(\lambda_1(\mathbf{\Gamma}) + \sigma_j^2(\tilde{\mathbf{G}})\right)$, and the Lagrange multiplier $\tilde{\mu}$ which satisfies the power constraint with equality is determined as

$$\tilde{\mu} = \frac{P_T + \sum_{i=1}^{L}\sum_{j=1}^{m}\frac{\text{Tr}(\mathbf{\Psi}_i)[\mathbf{\Phi}_i]_{jj}\lambda_j(\mathbf{R}_n)\lambda_1(\mathbf{\Gamma})}{\alpha_j}}{\sum_{j=1}^{m}\sqrt{\frac{\sum_{i=1}^{L}\text{Tr}(\mathbf{\Psi}_i)[\mathbf{\Phi}_i]_{jj}\lambda_j(\mathbf{R}_n)\sigma_j^2(\tilde{\mathbf{G}})}{\alpha_j^2}}}. \quad (26)$$

The values $p(j)$ upon substitution in the expression for $\tilde{\mathbf{\Omega}}$ in (20) yield $\tilde{\mathbf{F}}$. The individual precoders $\mathbf{B}_i$ can be determined as $\mathbf{B}_i = \mathbf{H}_i^\dagger\tilde{\mathbf{F}}_i$, where $\mathbf{H}_i^\dagger$ denotes the pseudo-inverse of $\mathbf{H}_i$. For the special case wherein the elements of the parameter



(a) (b)

Fig. 1: (a) MSE versus SNR_{FC} for ideal sensors (b) MSE performance of the proposed scheme with the existing schemes in [3] and [4] versus SNR_{FC} for ideal sensors.

vector $\boldsymbol{\theta}$ are uncorrelated, with $\mathbf{R}_\theta = \mathbf{I}_m$, the MSE of estimation in (18) reduces to

$$\text{MSE} = \text{Tr}\left(\left(\mathbf{I}_m + \tilde{\mathbf{G}}^H\tilde{\mathbf{F}}^H\left(\tilde{\mathbf{F}}\tilde{\mathbf{F}}^H + \mathbf{R}_n\right)^{-1}\tilde{\mathbf{F}}\tilde{\mathbf{G}}\right)^{-1}\right). \tag{27}$$

Upon substitution of the various quantities defined above, the MSE expression simplifies to

$$\text{MSE} = \sum_{j=1}^{m}\frac{p(j) + \lambda_j(\mathbf{R}_n)}{p(j)\left(1 + \sigma_j^2(\tilde{\mathbf{G}})\right) + \lambda_j(\mathbf{R}_n)}. \tag{28}$$

Solving the MSE optimization problem using the KKT framework, the closed form expression for $p(j)$ is obtained similar to (25) with $\lambda_1(\boldsymbol{\Gamma})$ replaced by 1. The optimal dual variable $\tilde{\mu}$ is given as

$$\tilde{\mu} = \frac{P_T + \sum_{j=1}^{m}\frac{1}{(1+\sigma_j^2(\tilde{\mathbf{G}}))}\sum_{i=1}^{L}\lambda_j(\mathbf{R}_n)\text{Tr}(\boldsymbol{\Psi}_i)[\boldsymbol{\Phi}_i]_{jj}}{\sum_{j=1}^{m}\frac{1}{(1+\sigma_j^2(\tilde{\mathbf{G}}))}\sqrt{\lambda_j(\mathbf{R}_n)\sigma_j^2(\tilde{\mathbf{G}})\sum_{i=1}^{L}\text{Tr}(\boldsymbol{\Psi}_i)[\boldsymbol{\Phi}_i]_{jj}}}. \tag{29}$$

A. BCRB and Centralized MMSE Bounds

To benchmark the MSE performance, the BCRB and centralized MMSE bounds are derived next. For the scenario with ideal sensors, the BCRB for the parameter $\boldsymbol{\theta}$, considering the observation model $\mathbf{y} = \mathbf{HF}\boldsymbol{\theta} + \mathbf{n}$, is obtained as [17]

$$\text{MSE}_{\text{BCRB}} \geq \text{Tr}\left(\left(\mathbf{R}_\theta{}^{-1} + \mathbf{F}^H\mathbf{H}^H\mathbf{R}_n^{-1}\mathbf{HF}\right)^{-1}\right). \tag{30}$$

For the wireless network with non-ideal sensors, the best performance is achieved when all the sensor observations are directly available at the fusion center. This is termed as the centralized MMSE benchmark. The corresponding concatenated observation vector at the fusion center is given as $\mathbf{x} = \mathbf{G}\boldsymbol{\theta} + \mathbf{v}$. The centralized MMSE bound for this scenario can be readily determined as

$$\text{MSE}_{\text{MMSE}} = \text{Tr}\left(\left(\mathbf{R}_\theta^{-1} + \mathbf{G}^H\mathbf{R}_v^{-1}\mathbf{G}\right)^{-1}\right), \tag{31}$$

where the overall observation vector $\mathbf{x} = [\mathbf{x}_1^H, \mathbf{x}_2^H, \ldots, \mathbf{x}_L^H]^H \in \mathbb{C}^{l\times 1}$. Simulation results are presented next to demonstrate the performance of parameter estimation and verify the various analytical formulations.

V. Simulation Results

This section presents simulation results to demonstrate the performance of the proposed schemes. For the simulation study, the elements of the channel and observation matrices $\mathbf{H}_i$ and $\mathbf{G}_i$, respectively, are generated as i.i.d. complex Gaussian random variables with zero mean and variance equal to unity. The parameter covariance matrix is set as $\mathbf{R}_\theta = \mathbf{I}_m$, unless otherwise mentioned explicitly, with $m = 3$. The number of transmit and receive antennas at each sensor i and the fusion center are set as $N_t = 3$ and $N_r = 3$, respectively, with the number of observations $l_i = 3$. The observation noise and channel noise covariance matrices are $\sigma_v^2\mathbf{I}_{N_t}$ and $\sigma_n^2\mathbf{I}_{N_r}$ respectively, with the SNR at each sensor defined as $\text{SNR}_v = \frac{1}{\sigma_v^2}$ and the fusion center SNR defined as $\text{SNR}_{\text{FC}} = \frac{1}{\sigma_n^2}$. The value of SNR_v is considered to be 20 dB.

Fig. 1(a) shows the plots of the resulting MSE for the precoders and combiner determined using the framework described in Section III for ideal sensors, versus the SNR_{FC} for different values of the number of sensors $L \in \{5, 10, 20\}$ in the network. It is seen that MSE monotonically decreases and also coincides with the BCRB derived in (30) for the above setup, which demonstrates the efficiency of the proposed scheme. Moreover, as the number of sensors increases, the MSE is seen to progressively decrease, as can be naturally expected due to the availability of an increasing number of observations at the fusion center. Fig. 1(b) shows the MSE performance of

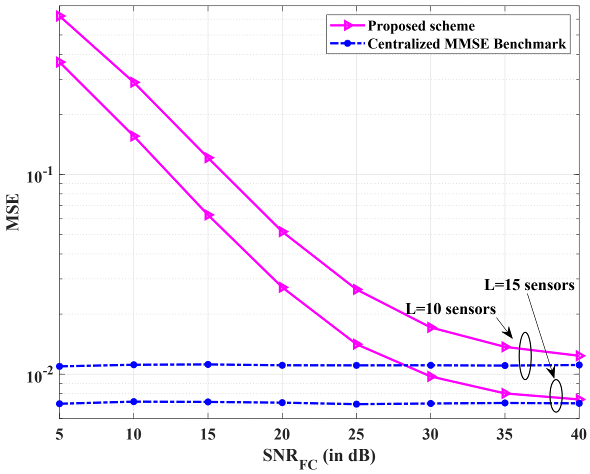


Fig. 2: MSE versus SNR_{FC} for non-ideal sensors

the scheme proposed in Section III and the existing schemes in [3] and [4]. The total number of sensors in the WSN are set equal to 10. It can be seen that the proposed design obtains a lower MSE in comparison to the existing iterative schemes due to the optimality of the proposed design.

Fig. 2 depicts the MSE performance of the scheme proposed in Section IV for the scenario with non-ideal sensors, against the SNR at the fusion center. It can once again be seen that the MSE decreases as the number of sensors increases. Also, the proposed scheme attains the centralized MMSE bound at high SNR.

VI. Conclusion

This paper examined the problem of linear decentralized estimation of an unknown vector parameter in a MIMO sensor network over a coherent MAC. The precoder and combiner designs are based on majorization theory and, therefore are, non-iterative in nature, in contrast to competing schemes in the existing literature. The decentralized estimation framework was presented initially for a scenario with ideal sensors. This was subsequently extended to non-ideal sensors, considering also a general scenario with inter parameter correlation. Performance benchmarks such as the BCRB and the centralized MMSE bound were derived to characterize the best achievable MSE performance for linear estimation in the sensor network. Simulation results demonstrated the performance of the proposed schemes and also the achievement of the respective bounds.

Appendix A
Optimality of Precoder Design

A real valued function f_0 defined on a set $\mathcal{A} \subseteq \mathbb{R}^n$ is Schur concave on $\mathcal{A}$ if

$$\mathbf{x} \prec \mathbf{y} \text{ on } \mathcal{A} \Rightarrow f_0(\mathbf{x}) \geq f_0(\mathbf{y}).$$

Using the result from majorization theory [18], for a Hermitian matrix $\mathbf{R}$, $\mathbf{d}(\mathbf{R}) \prec \boldsymbol{\lambda}(\mathbf{R})$, where $\mathbf{d}(\mathbf{R})$, $\boldsymbol{\lambda}(\mathbf{R})$ denote the vectors comprising of the principal diagonal elements and the eigenvalues of the matrix $\mathbf{R}$, respectively. From the above definition of a Schur concave function, it follows that

$$f_0(\mathbf{d}(\mathbf{E})) \geq f_0(\boldsymbol{\lambda}(\mathbf{E})), \tag{32}$$

for any arbitrary error covariance matrix $\mathbf{E}$. Since the MSE cost function chosen in this paper is Schur concave in nature [12], the lower bound on the MSE is achieved when $\mathbf{d}(\mathbf{E}) = \boldsymbol{\lambda}(\mathbf{E})$, which holds true for a diagonal matrix. Therefore, the proposed schemes in Sections III for the ideal noiseless sensors achieve the lower bound on the MSE, since the effective error covariance matrix $\tilde{\mathbf{E}}$ is diagonalized. Hence, the precoders designed is optimal.

References

[1] J. Yick, B. Mukherjee, and D. Ghosal, "Wireless sensor network survey," *Computer networks*, vol. 52, no. 12, pp. 2292–2330, 2008.

[2] J. Xiao, S. Cui, Z. Luo, and A. J. Goldsmith, "Linear coherent decentralized estimation," *IEEE Transactions on Signal Processing*, vol. 56, no. 2, pp. 757–770, 2008.

[3] A. S. Behbahani, A. M. Eltawil, and H. Jafarkhani, "Linear decentralized estimation of correlated data for power-constrained wireless sensor networks," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6003–6016, 2012.

[4] Y. Liu, J. Li, and X. Lu, "Joint transceiver design for linear MMSE data fusion in coherent MAC wireless sensor networks," *Information Fusion*, vol. 37, pp. 37–49, 2017.

[5] N. K. Venkategowda and A. K. Jagannatham, "Optimal minimum variance distortionless precoding (MVDP) for decentralized estimation in MIMO wireless sensor networks," *IEEE Signal Processing Letters*, vol. 22, no. 6, pp. 696–700, 2015.

[6] F. Jiang, J. Chen, and A. L. Swindlehurst, "Optimal power allocation for parameter tracking in a distributed amplify-and-forward sensor network," *IEEE Transactions on Signal Processing*, vol. 62, no. 9, pp. 2200–2211, May 2014.

[7] J. Akhtar and K. Rajawat, "Distributed sequential estimation in wireless sensor networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 1, pp. 86–100, Jan 2018.

[8] A. Shirazinia, S. Dey, D. Ciuonzo, and P. Salvo Rossi, "Massive MIMO for decentralized estimation of a correlated source," *IEEE Transactions on Signal Processing*, vol. 64, no. 10, pp. 2499–2512, May 2016.

[9] S. Kar and P. K. Varshney, "Linear coherent estimation with spatial collaboration," *IEEE Transactions on Information Theory*, vol. 59, no. 6, pp. 3532–3553, June 2013.

[10] S. Liu, S. Kar, M. Fardad, and P. K. Varshney, "Sparsity-aware sensor collaboration for linear coherent estimation," *IEEE Transactions on Signal Processing*, vol. 63, no. 10, pp. 2582–2596, May 2015.

[11] ——, "Optimized sensor collaboration for estimation of temporally correlated parameters," *IEEE Transactions on Signal Processing*, vol. 64, no. 24, pp. 6613–6626, Dec 2016.

[12] D. P. Palomar and Y. Jiang, "MIMO transceiver design via majorization theory," *Foundations and Trends® in Communications and Information Theory*, vol. 3, no. 4-5, pp. 331–551, 2007.

[13] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1993.

[14] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[15] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.

[16] R. A. Horn and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1990.

[17] H. L. V. Trees and K. L. Bell, *Bayesian Bounds for Parameter Estimation and Nonlinear Filtering/Tracking*. Hoboken, NJ, USA: Wiley, 2007.

[18] A. W. Marshall, I. Olkin, and B. C. Arnold, *Inequalities: theory of majorization and its applications*. Springer, 1979, vol. 143.